

INTERVIEW

»Der Earth Simulator hat uns aufgeweckt«

Der Chef des kalifornischen Hochleistungs-Rechenzentrums NERSC äußert sich zu der Herausforderung durch den jüngsten japanischen Supercomputer.

Spektrum: Im April 2002 wurde in Japan der »Earth Simulator« in Betrieb genommen, der den bis dahin schnellsten Rechner der Welt, »ASCI White«, um das Fünffache an Geschwindigkeit übertraf (Spektrum der Wissenschaft 9/2002, S. 14). Damals hieß es, über die Amerikaner sei ein »Computenik-Schock« gekommen, vergleichbar dem Sputnik-Schock von 1957. Hat der Earth Simulator die Fachleute so überrascht?

Simon: Der Rechner selbst nicht. Die japanischen Kollegen hatten schon vor zwei Jahren, noch vor der offiziellen Ankündigung, mit uns über das Projekt gesprochen. Aber dass die Herstellerfirma NEC den Earth Simulator pünktlich zum Liefertermin nicht nur aufgestellt hatte, sondern er auch funktionierte, das war eine große Überraschung.

Man muss dazu wissen, dass sich unter den amerikanischen Herstellern eine »Kultur der Verspätung« etabliert hat: Größere Innovationen kommen regelmäßig drei bis sechs Monate nach dem angekündigten Termin. Von der für 2002 angekündigten amerikanischen Spitzenmaschine, »ASCI-Q«, die mit 30 Teraflops Spitzenleistung in Los Alamos arbeiten sollte, sind bis heute nur zwei 10-Teraflops-Rechner in Betrieb.

Die zweite Überraschung: Nach unseren Gesprächen hatte ich erwartet, dass der Earth Simulator auf eine Dauerleistung von vielleicht zehn Teraflops kommen würde. Tatsächlich hat er, mit einer Anwendung aus der Klimaforschung, 27 von 40 möglichen Teraflops erreicht. Das ist sowohl absolut als auch im Wirkungsgrad eine überragende Leistung; sie wurde im letzten Herbst mit dem Gordon-Bell-Preis ausgezeichnet.

Spektrum: Das Konstruktionsprinzip für die Einzelrechner des Earth Simulator, der Vektorrechner, gilt außerhalb Japans weitgehend als überholt. Ist der Earth Simulator so erfolgreich, weil oder obwohl er aus Vektorrechnern besteht?

Simon: Weder noch. Das für den Erfolg entscheidende Konstruktionsmerkmal ist die große Bandbreite zwischen Speicher und Prozessor [siehe Glossar]. Gerade für Klimasimulationen ist es typisch, dass der Prozessor eine Zahl aus dem Speicher holt, mit ihr eine einzige Multiplikation durchführt und das Ergebnis gleich wieder abspeichert. Wenn das Verhältnis von Rechenoperationen zu Speicherzugriffen, wie in diesem Fall, in die Nähe von 1 gerät, muss die Architektur sehr gut ausgewogen sein, damit die Maschine auf einen guten Wirkungsgrad kommt.

Spektrum: Gibt es eine Hauptursache für den Erfolg?

Simon: Die Bereitschaft, viel Geld aufzuwenden! Bandbreite für Speicherzugriff ist einfach sehr teuer. Aber wenn in den USA die Regierung bereit ist, ebenfalls 400 Millionen Dollar für einen Superrechner zu bezahlen, dann können wir das auch. Die »Cray X-1« hat eine sehr gute Bandbreite, kostet aber auch entsprechend.



Horst Simon hat an der Technischen Universität Berlin 1978 das Diplom in Mathematik abgelegt und 1982 an der Universität von Kalifornien in Berkeley promoviert. Seit 1996 ist er Direktor des National Energy Research Scientific Computing Center (NERSC), das zum Lawrence-Berkeley-Nationallaboratorium in Berkeley (Kalifornien) gehört. Als überregionales Hochleistungsrechenzentrum ist das NERSC deutschen Zentren wie dem Konrad-Zuse-Institut in Berlin oder dem John-von-Neumann-Institut für Computing in Jülich vergleichbar.

Spektrum: Lohnt das den Aufwand?

Simon: Nicht unter allen Umständen. Es macht keinen Sinn, für eine Verdopplung des Wirkungsgrades den fünffachen Preis zu bezahlen. Andererseits ist der Wirkungsgrad eine sehr schlecht bestimmte Größe, weil er stark von der Art der Anwendung abhängt. Es gibt Programme, die laufen auf einer »SX-6« – dem Grundbaustein des Earth Simulator – mit 60 Prozent und nur mit fünf bis zehn Prozent auf einer »Power 3« oder »Power 4« von IBM. Bei anderen Programmen ist es genau umgekehrt.

Spektrum: Was kommt nach ASCI-Q?

Simon: Die Kollegen vom Lawrence-Livermore-Nationallaboratorium haben schon »ASCI Purple« angekündigt. Es wird sich um eine IBM-Maschine handeln, die das Stadium Power 4 überspringt und gleich mit Power-5-Prozessoren ausgestattet sein wird. Die Auslieferung ist für Ende 2004 geplant, die genaue Leistung steht noch nicht fest. Unter Vorbehalt sind 100 Teraflops angekündigt.

Spektrum: Gibt es andere Neuentwicklungen in den USA?

Simon: Ja. Wir kommen allmählich von einer strikten Entweder-Oder-Haltung ab, die etwa besagt: Entweder besteht ein Superrechner ganz aus relativ billigen Prozessoren aus der Massenfertigung, und man gleicht ihre mangelhafte Eignung durch große Anzahlen aus – das sind die Clusters und die meisten Geräte, die IBM anbietet; oder sämtliche Teile des Rechners werden für diesen Zweck neu entworfen – Beispiel Cray X-1 –, was wegen der geringen Stückzahl die einzelne Maschine jedoch sehr teuer macht. Inzwischen sind sehr interessante Mischungen dieser Extreme in der Planung.

Das Projekt »Red Storm« ist eine solche Mischung aus 80 Prozent Standardware und entscheidenden 20 Prozent Spezialentwicklung. Die Sandia Laboratories in Albuquerque (New

Mexico) haben die Architektur dieses Rechners selbst konzipiert und die Herstellung bei Cray in Auftrag gegeben.

Das Projekt »Blue Planet«, das wir bei NERSC in Zusammenarbeit mit IBM entwickelt haben, wird eine ähnliche Mischung realisieren. IBM hatte schon entschieden, wie ihre kommerziellen Maschinen 2005/2006 aussehen sollen, und wir haben Veränderungen vorgeschlagen, welche die Leistung für wissenschaftliche Zwecke erheblich erhöhen würden.

Spektrum: Wird denn »Blue Planet« an Ihrem Zentrum stehen?

Simon: Das hängt von den Haushaltsmitteln ab. Wir sind sehr überzeugt von unserem Projekt und würden sogar die Prognose wagen, dass das Gerät die doppelte Leistung des Earth Simulator zum halben Preis bietet. Aber das wären immer noch 200 Millionen Dollar. Für ein ziviles Zentrum mit Grundlagenforschung wie unseres galt bisher eher ein Zehntel dieses Betrages als angemessen.

Beide Häuser des Kongresses sind der Wissenschaft im Allgemeinen und der Grundlagenforschung im Besonderen sehr wohlgesonnen. Man versteht mittlerweile, dass nicht nur Biologie und Medizin, sondern auch Physik, Mathematik und Informatik sehr wichtig sind. Andererseits hat die Regierung Steuerersenkungen versprochen, und der Irak-Krieg wird den Staatshaushalt sehr belasten.

Spektrum: Also gibt es keine Antwort der USA auf den Earth Simulator?

Simon: Nicht in diesem und nicht im nächsten Jahr. Im Haushaltsentwurf für 2004 ist keine nennenswerte Ausgabensteigerung

für Hochleistungsrechner vorgesehen. Immerhin hat der Earth Simulator die Wissenschaftler und Politiker aufgeweckt.

Spektrum: Was bedeutet das?

Simon: Vor etwa zehn Jahren war die japanische Computerindustrie noch weit bedeutender als heute. Es gab kleinere Hersteller von Höchstleistungsrechnern in Großbritannien und in Deutschland. Um diese Zeit wurde die Initiative HPCC (*high performance computing and communications*) gestartet, mit dem Ergebnis, dass die Dominanz der USA in Hardware wie in Software sehr stark ausgebaut wurde.

Seit etwa sechs Jahren hat diese Initiative keinen neuen An Schub mehr enthalten. Ein Petaflop-Workshop führte zu nichts, weil es keine neuen Fördermittel gab. Aus diesem Dämmer schlaf sind wir aufgeweckt worden. Es gibt wieder gemeinsame Ausschüsse (*task forces*) der verschiedenen Wissenschaftsförderungs-Agenturen. Die National Academy of Sciences hat eine Studie über die Zukunft des Hochleistungsrechnens in Auftrag gegeben. Es gärt so wie seit acht Jahren nicht mehr.

Spektrum: In der Liste der 500 stärksten Rechner finden sich immer mehr Cluster, also »Haufen« gewöhnlicher PCs, die über Datenleitungen verbunden sind und sich eine große Rechenarbeit teilen. Wie kommt es, dass ausgerechnet der Pionier Thomas Sterling sein eigenes Kind, den Beowulf-Cluster (Spektrum der Wissenschaft 3/2002, S. 88), für tot erklärt?

Simon: Cluster sind nicht wirklich tot. Sie sind geeignet für den Rechenbedarf einer einzelnen Abteilung mit hundert Leuten in der Industrie, oder für einen Universitäts-Fachbereich. Wir haben im Forschungszentrum zehn bis fünfzehn PC-Cluster aus 32 bis 512 Prozessoren. Die sind sehr nützlich, zum Üben für die Wissenschaftler, bevor sie auf das große Gerät gehen.

Thomas Sterling hat zwei Kinder. Der PC-Cluster Beowulf ist das eine. Das andere ist HTMT (*hybrid technology multithreaded system*, Spektrum der Wissenschaft 1/2002, S. 78), und das war der einzige ernstzunehmende Bauprinzip, das vor 2008 eine Leistung im Petaflop-Bereich hätte erbringen können. Ein solches Ziel mit PC-Clustern anzustreben ist aussichtslos oder zumindest vom Aufwand her ziemlich abwegig. Das hat Thomas Sterling gemeint, als er im letzten Jahr in Baltimore Beowulf für tot erklärte.

Spektrum: Kann das Grid den Rechenbedarf der Kunden Ihres Zentrums befriedigen?

Simon: PC-Cluster und Grid Computing werden häufig in einem Atemzug genannt, sind aber grundsätzlich verschiedene Dinge. Das Grid wird ja gerne mit dem Stromnetz verglichen. Und zu den Leuten, die glauben, das Grid würde das Hochleistungsrechnen überflüssig machen, kommt mir der alte Spottspruch der Atomindustrie in den Sinn: »Wozu Kraftwerke? Bei uns kommt der Strom aus der Steckdose.« Supercomputer sind so etwas wie die Kraftwerke für das Grid.

Spektrum: Und dass der kleine PC-Benutzer seine Rechenleistung mit ins Grid einspeist?

Simon: Das mag für manche Anwendungen nützlich sein; für echtes Hochleistungsrechnen ist es uninteressant.

Das Interview führte **Christoph Pöppe**, Redakteur bei Spektrum der Wissenschaft.

GLOSSAR

- ▶ Die **Leistung eines Supercomputers** wird in »Flops« (*floating point operations per second*, Gleitkomma-Rechenakte pro Sekunde) gemessen. Der zurzeit schnellste Rechner der Welt, der **Earth Simulator**, hat eine theoretische Spitzenleistung (*peak performance*) von 40 Teraflops (Billionen Flops); diese würde erreicht, wenn alle in ihm enthaltenen Einzelrechner (Prozessoren) gleichzeitig aktiv wären. Das ist jedoch praktisch nie der Fall, weil die Prozessoren häufig auf Daten warten müssen. Den Anwender interessiert weit mehr die bei der Ausführung seines Programms erbrachte **Dauerleistung** (*sustained performance*).
- ▶ Ein **Petaflops** sind 1000 Teraflops oder 10^{15} Flops.
- ▶ Eine Messgröße für die Effizienz eines Superrechners – unabhängig von dessen schierer Größe – ist der **Wirkungsgrad**, das Verhältnis von Dauer- zu Spitzenleistung (*sustained to peak performance ratio*).
- ▶ **Bandbreite** (genauer: *processor to memory bandwidth*) ist »die Breite des Kanals, auf dem die Daten vom Arbeitsspeicher zum Rechenwerk schwimmen«, das heißt die Anzahl an Dateneinheiten, die pro Sekunde übermittelt werden können.
- ▶ **Power 3** und **Power 4** heißen die Prozessoren, die IBM gegenwärtig in ihre Superrechner einbaut. Power 3 wird seit 2000 ausgeliefert, Power 4 seit 2002.